AD-A070 882   MINNESOTA UNIV ST PAUL DEPT OF APPLIED STATISTICS   F/G 12/1
ALTERNATIVE COMPUTATIONAL METHODS FOR ESTIMATION OF LOGIT MODEL--ETC(U)
JUN 79   S E FIENBERG, M M MEYER, G W STEWART   N00014-78-C-0600

UNCLASSIFIED   NL
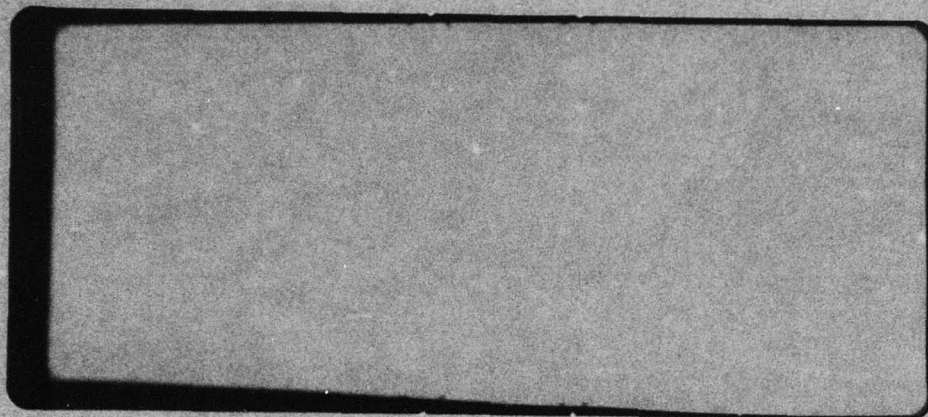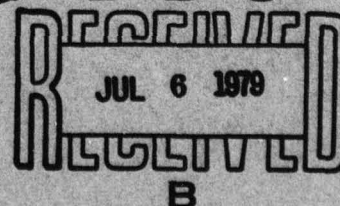
AD
A070882

END
DATE
FILMED

8-79

DDC

12 VSS. LEVEL II

See 1473 in back.

UNIVERSITY OF MINNESOTA

SCHOOL OF STATISTICS

D D C
RECEIVED
JUL 6 1979
B

79 07 06 018

ALTERNATIVE COMPUTATIONAL METHODS FOR
ESTIMATION OF LOGIT MODEL PARAMETERS

by

Stephen E. Fienberg and Michael M. Meyer
University of Minnesota
and
G.W. Stewart
University of Maryland

Technical Report No. 348

June 1979

ALTERNATIVE COMPUTATIONAL METHODS FOR
ESTIMATION OF LOGIT MODEL PARAMETERS

Stephen E. Fienberg and Michael M. Meyer
School of Statistics, University of Minnesota
St. Paul, Minnesota, U.S.A.

and

G. W. Stewart
Department of Computer Science, University of Maryland
College Park, Maryland, U.S.A.

INTRODUCTION The analysis of cross-classified categorical data involves statistical problems where both the explanatory variables (or factors) and response variables are categorical. Loglinear and logit models are now widely-used in the analysis of such data (e.g. see Bishop, Fienberg, and Holland, 1975; Fienberg 1977; Haberman, 1974, 1978). Since logit models are loglinear models, the computational methods for the analysis of cross-classifications via loglinear models can also be used for analyses involving the use of logit models.

In this paper we compare three different computational approaches for maximum likelihood estimation in logit situations:
- (a) iterative proportional fitting,
- (b) iteratively reweighted least squares as implemented in GLIM (see Nelder and Wedderburn, 1972),
- (c) a variant of Newton's method, as developed by Fienberg and Stewart (1979), applied in a somewhat different form for loglinear and logit formulations.

Additional comparisons can be made with the Newton-Raphson algorithm of Haberman (1978), but they are not included here.

LOGLINEAR AND LOGIT MODELS For simplicity, consider a problem involving two explanatory variables with dimensions I and J , and a response variable of dimension K . Thus the data are counts in the form of an I×J×K table where the totals in the I×J margin corresponding to the explanatory variables are taken as fixed. We assume that the sampling model for the counts is product-multinomial (Bishop, Fienberg, and Holland, 1975). Logit models (involving K-1 simultaneous logit equations) for this problem are equivalent to loglinear models that treat the three variables as if they are responses, and that include u-terms corresponding to the main effects and interaction for the explanatory variables (Fienberg, 1977, Chapter 6). Thus the iterative proportional fitting (IPF) algorithm used for loglinear models, can be used directly for logit models in this problem.

NEWTON'S METHOD Fienberg and Stewart (1979) have used a variant of
Newton's method for analyzing both loglinear and logit models. Their
first algorithm treats the logit model parameters as loglinear model
parameters, and adjusts for the required conditioning at the end of
the computation. A second algorithm proceeds by initially conditioning on
the explanatory variables, and then using a somewhat different set
of computations.

These algorithms involve the construction of the upper half of
a $p \times p$ weighted cross-product matrix, and take full advantage
of the sparseness of the $n \times p$ design matrix without actually con-
structing it. The algorithms proceed via Newton's method with
variable step length, using a Cholesky decomposition with pivoting.
Further details will be reported in the near future.

GLIM The GLIM algorithm, as developed by Nelder and his colleagues,
is designed to handle the product multinomial sampling model for
binary response structures, as well as other sampling models not
considered here. For further details see Nelder and Wedderburn
(1972). In order to handle a K-level response variable in GLIM, the
user needs to treat it in an asymmetric fashion, eg. via the use of
continuation ratios (see Fienberg, 1977, Chapter 6).

COMPARISONS We have *summarized the qualitative aspects* of the four
algorithms in question in Table 1. Because the Fienberg-Stewart
algorithms have opted for economy of storage over efficiency of
operation, the comparisons of storage requirements between their
algorithms and GLIM is misleading. In practice GLIM cannot handle
as large problems as can be handled by Fienberg-Stewart algorithms.

Although we have not made direct comparisons here on speed of
convergence, we note that IPF has linear convergence properties
while the other algorithms have quadratic convergence. We expect
that the special features in the Fienberg-Stewart algorithms should
allow for convergence at a slightly faster rate than GLIM, but this
should not be a serious difference between these methods. More
important is the issue of numerical stability of the algorithms,
where again we anticipate the superiority of the Fienberg-Stewart
ones.

Whether one should use the logit or loglinear version of the
Fienberg-Stewart algorithms depends on the size of the marginal
array corresponding to the explanatory variables. When this margin
is small some advantages may accrue to the loglinear approach.

REFERENCES

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). _Discrete Multivariate Analysis: Theory and Practice_. Cambridge, Mass., The MIT Press.

Fienberg, S.E. (1977). _The Analysis of Cross-classified Categorical Data_. Cambridge, Mass., The MIT Press.

Fienberg, S.E. and Stewart, G.W. (1979). "The numerical analysis of contingency tables." Unpublished manuscript.

Haberman, S.J. (1974). _The Analysis of Frequency Data_. Chicago, Ill., University of Chicago Press.

Haberman, S.J. (1978). _Analysis of Qualitative Data, Volume I, Intro-ductory Topics_. New York, Academic Press.

Nelder, J.A. and Wedderburn, R.W.M. (1972). Generalized linear models. _J. Roy. Statist. Soc. A_ 135, 370-384.

SUMMARY Several algorithms have been proposed for the computation of maximum likelihood estimates for contingency tables. Since logit models can be treated as special versions of loglinear models, many of these same techniques can be used for logit models as well. In this paper, we compare in a qualitative fashion the relative merits of (i) the widely-used method of iterative proportional fitting, (ii) GLIM as developed by Nelder and Wedderburn, and (iii) two variants of Newton's method developed by Fienberg and Stewart.

SOMMAIRE Plusieurs algorithmes ont été proposés afin de calculer les évaluations de probabilité maximum pour les tables d'éventualité. Puisque les modèles "logit" peuvent être traités comme un cas parti-culier des modèles logarithmiques et linéaires, la plupart des tech-niques qui s'appliquent à ces derniers peuvent s'appliquer dont aussi bien aux modèles "logit". Dans cet article, nous comparons de facon qualitative les mérits relatifs des méthodes suivantes: (i) la méthode fréquemment utilisée de l'ajustement proportionnel itératif, (ii) la méthode GLIM, développée par Nelder et Wedderburn, et (iii) la méthode des deux variants de Newton, développée par Fienberg et Stewart.

Table 1: General Comparison of Methods.

Criterion	(i) Fienberg-Stewart Loglinear	(ii) Fienberg-Stewart Logit	(iii) GLIM	(iv) IPF
1. Utilizes conditioning implied by logit model	No	Yes	Yes	No
2. Handles Polytomous response structures directly	Yes	Yes	No	Yes
3. Can easily be extended to general logistic regression	No	Yes	Yes	No
4. Storage requirements	SSP matrix[1].	SSP matrix[2].	SSP matrix[2].	Data array, plus margins.
5. Detects non-existence of MLE's	Yes	Yes	No, only by non-convergence.	No, only by slow convergence to zero.
6. Handles structural zeros	Yes	Yes	No	Yes
7. Produces parameter estimates	Yes	Yes	Yes	Only for complete tables
8. Produces Covariance estimates	Yes	Yes	Yes	No
9. Detects Aliasing	Yes, via pivoted Cholesky decomposition.	Yes, via pivoted Cholesky decomposition.	Yes, via small diagonal elements in non-pivoting Gaussian elimination.	Not applicable.

1 SSP matrix is for all loglinear parameters.

2 SSP matrix is for logit parameters only.

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 348 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER

1. REPORT NUMBER Technical Report No. 348

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

ALTERNATIVE COMPUTATIONAL METHODS FOR ESTIMATION OF LOGIT MODEL PARAMETERS

5. TYPE OF REPORT & PERIOD COVERED

Technical rept.

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Stephen E. Fienberg, Michael M. Meyer, G.W. Stewart

8. CONTRACT OR GRANT NUMBER(s)

N00014-78-C-0600

N00014-78-C-0751

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Department of Applied Statistics.
School of Statistics, University of Minnesota
1994 Buford Avenue, St. Paul, MN 55108

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

12. REPORT DATE

June 1979

13. NUMBER OF PAGES

4

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

7 p.

15. SECURITY CLASS. (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Categorical data; logit models; loglinear models; iterative
proportional fitting; GLIM; Newton's method.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we compare three different computational approaches
for maximum likelihood estimation in logit situations:
(a) iterative proportional fitting,
(b) iteratively reweighted least squares as implemented in GLIM
(see Nelder and Wedderburn, 1972),
(c) a variant of Newton's method, as developed by Fienberg and
Stewart (1979), applied in a somewhat different form for loglinear
and logit formulations.